MEME16203Linear Models

Assignment 2

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	May 2022		
Due by:	27/6/2022		

Q1. Suppose that we are interested in the coefficients $\boldsymbol{\beta}$ of a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y}$ is $n \times 1$, $\mathbf{X}$ is nonsingular with dimension $n \times p$ and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$. Furthermore, suppose that it is of interest to partition that model in the form $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3\end{bmatrix}$, for $n \times p_i$ matrices $\mathbf{X}_i, i = 1, 2, 3$. Finally, suppose that an investigator creates a partially orthogonal design, in which $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3\end{bmatrix}$ has the property that $\mathbf{X_i^T X_j} = 0$ for $i \neq j$. We have been showing that the least squares estimate of $\boldsymbol{\beta}$ takes the form $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3\end{bmatrix}$, where

- $\hat{\boldsymbol{\beta}}_\mathbf{1} = (\mathbf{X_1^T X_1})^{-\mathbf{1}}\mathbf{X_1^T Y}$
- $\hat{\boldsymbol{\beta}}_\mathbf{2} = (\mathbf{X_2^T X_2})^{-\mathbf{1}}\mathbf{X_2^T Y}$
- $\hat{\boldsymbol{\beta}}_\mathbf{3} = (\mathbf{X_3^T X_3})^{-\mathbf{1}}\mathbf{X_3^T Y}$

Show that the estimates $\hat{\boldsymbol{\beta}}_\mathbf{1}$, $\hat{\boldsymbol{\beta}}_\mathbf{2}$ and $\hat{\boldsymbol{\beta}}_\mathbf{3}$ are uncorrelated. That is, show that the covariance matrix $V(\hat{\boldsymbol{\beta}})$ has a block-diagonal form

$$\begin{bmatrix}\mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_3\end{bmatrix}$$

for some matrices $\mathbf{A}_1$,$\mathbf{A}_2$ and $\mathbf{A}_3$. (10 marks)

Q2. You are given:

- Model (1): $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$
- Model (2): $Y_{ij} = \gamma_0 + \gamma_1(X_i - 10) + \gamma_2(X_i - 10)^2 + \epsilon_{ij}$

where $i = 1, 2, 3,\ j = 1, 2,\ X_1 = 5, X_2 = 10, X_3 = 15$ and μ, α_1, α_2, α_3, γ_0, γ_1, γ_2 and γ_3 are unknown parameters.

(a) For model (2), write down a formula for the best linear unbiased estimator (BLUE) for $\boldsymbol{\gamma} = \begin{bmatrix}\gamma_0 \\ \gamma_1 \\ \gamma_2\end{bmatrix}$. (10 marks)

(b) For model (1), verify that $\tau = 3\mu + 3\alpha_1 - 6\alpha_2 + 6\alpha_3$ is an estimable function and writhe down a formula for $\hat{\tau}$, the BLUE for τ. (10 marks)

(c) Formulate what is meant by the statement that model (2) is a reparameterization of model (1), and verify that this statement is correct. (10 marks)

Q3. Suppose that y_{11} and y_{12} are independent $N(\mu_1, 9\sigma^2)$ variables independent of y_{21} and y_{22} that are independent $N(\mu_2, 16\sigma^2)$ and $N(\mu_2, 16\sigma^2)$ variables respectively. What is the BLUE of $4\mu_1 + 2\mu_2$? Explain carefully. (20 marks)

Q4. Consider a problem of quadratic regression in one variable, $\mathbf{X}$. In particular, suppose that $n = 5$ values of a response $\mathbf{y}$ are related to values $x = 0, 1, 2, 3, 4$ by a linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ for

$$\mathbf{y} = \begin{bmatrix} 2 \\ 6 \\ 6 \\ 11 \\ 12 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \text{ and } \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

Define

$$\mathbf{W} = \begin{bmatrix} 1 & -2 & 2 \\ 1 & -1 & -1 \\ 1 & 0 & -2 \\ 1 & 1 & -1 \\ 1 & 2 & 2 \end{bmatrix}$$

Show that $\mathbf{y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ is reparameterization of $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\gamma}^{\boldsymbol{T}} = [\gamma_1, \gamma_2, \gamma_3]$. Find the OLS estimate of $\boldsymbol{\gamma}$ in the model $\mathbf{y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ and then OLS estimate of $\boldsymbol{\beta}$ in the original model. (Find numerical values.) (20 marks)

Q5. Suppose $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where for $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)'$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$. A particular experiment produces $n = 5$ data points as per

x_i	20	39	28	37	85
y	305	363	124	140	89

Suppose that $V(\boldsymbol{\epsilon}) = \sigma^2 diag\left(\frac{1}{4}, \frac{1}{16}, \frac{1}{16}, \frac{1}{49}, \frac{1}{100}\right)$. Evaluate an appropriate BLUE of $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ under the model assumptions. (20 marks)